The extent and drivers of gender imbalance in neuroscience reference lists

Jordan D. Dworkin[1], **Kristin A. Linn**[1], **Erin G. Teich**[2], **Perry Zurn**[3], **Russell T. Shinohara**[1], **and Danielle S. Bassett**[2,4,5,6,7,8,✉]

[1]Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics,Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[2]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Philosophy and Religion, American University, Washington DC, USA
[4]Department of Physics & Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA
[5]Department of Electrical & Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA
[6]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[7]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[8]Santa Fe Institute, Santa Fe, NM, USA

Like many scientific disciplines, neuroscience has increasingly attempted to confront pervasive gender imbalances within the field. While much of the conversation has centered around publishing and conference participation, recent research in other fields has called attention to the prevalence of gender bias in citation practices. Because of the downstream effects that citations can have on visibility and career advancement, understanding and eliminating gender bias in citation practices is vital for addressing inequity in a scientific community. In this study, we sought to determine whether there is evidence of gender bias in the citation practices of neuroscientists. Utilizing data from five top neuroscience journals, we indeed find that reference lists tend to include more papers with men as first and last author than would be expected if gender was not a factor in referencing. Importantly, we show that this overcitation of men and undercitation of women is driven largely by the citation practices of men, and is increasing with time despite greater diversity in the academy. We develop a co-authorship network to determine the degree to which homophily in researchers' social networks explains gendered citation practices and we find that men tend to overcite other men even when their social networks are representative of the field. We discuss possible mechanisms and consider how individual researchers might incorporate these findings into their own referencing practices.

Neuroscience | Bibliometrics | Gender
Correspondence: ***dsb@seas.upenn.edu***

Introduction

In recent years, science has been pushed to grapple with the social and structural systems that produce vast gender imbalances in academic participation. Research has found large and persistent gaps in the proportion of women across scientific fields and has estimated that many fields will not reach gender equity for decades at their current trajectories (1). For women currently or formerly in academia, gender imbalances have persisted across various measures of academic inclusion and success. Prior work has found that such biases are present in compensation (2), grant funding (3–5), credit for collaborative work (6), teaching evaluations (7–9), hiring and promotions (10–12), and number of papers published and cited (13–17). Importantly, while this study focuses on gender, similar biases have been demonstrated in domains like race, socioeconomic status, and university prestige (18, 19).

While many aspects of gender bias have yet to be studied within neuroscience specifically, issues of gender and diversity have commanded increasing attention over the past several years. Groups like BiasWatchNeuro (biaswatchneuro.com), Women in Neuroscience (winrepo.org), and Anne's List (anneslist.net) have been created to track and promote the inclusion of women in conferences and symposia. Furthermore, major neuroscience societies have publicly discussed ways to improve representation (20), and journals have sought to balance the composition of editors and reviewers (21). On the heels of these efforts, a recent study showed that authorship and public speaking have indeed become more balanced in the last decade (22).

However, measures of authorship and conference participation reflect only one aspect of success in a field, and the presence of differential engagement with scholarship could lead to prolonged inequities in other areas. Recent studies of such differential engagement have found not only that people from marginalized groups are broadly undercited in fields such as communications (23) and philosophy (24), but also that women-led research in particular tends to receive fewer citations than comparable papers led by men in the fields of astronomy (17), international relations (16), and political science (25). Theoretical work has proposed a "Matilda effect" in which the contributions of men are seen as more central within a field and are therefore sought out more often and evaluated more highly (26). In visual art and literary texts, the "Bechdel test" has revealed the prevalence of cases in which women's contributions are not valued independently of men (27, 28). The presence of such an effect in scientific authorship would likely produce reputational and citational inequity. In this case, women-led work could remain under-discussed and perceived as more marginal than men-led work.

Because of the potential for harmful downstream effects of inequitable engagement with women and men's work, the study of citation behavior is a critical endeavor for understanding and addressing a field's biases. Additionally, achieving gender equity within citation lists is a goal that can be pursued by all researchers during their paper-writing process (unlike, for example, achieving gender equity within keynote speaker roles). Thus, in this study we seek to determine the existence and potential drivers of gender bias in neuroscience citations. Previous work in citation gaps has often focused on the relationship between authors' gender and their citation counts (16, 17), finding that work by women tends to receive fewer citations than similar work by men. Yet this formulation only measures the passive consequences of gendered citation behavior, rather than directly measuring the behavior itself. Instead, building on recent studies conducted in international relations and political science (25, 29), we investigate the relationship between authors' gender and the gender make-up of their reference lists. Using this framework, we are able to quantify properties associated with authors serving as both objects and agents of undercitation.

For this study, we examine the authors and reference lists of articles published in five top neuroscience journals since 1995. Within this pool of articles, we are able to obtain probabilistic estimates of authors' gender identity, find connections between citing and cited papers, locate and remove instances of self-citation, and study the links between authors' genders and their role as objects/agents of undercitation. Specifically, we test the following hypotheses: (1) The overall citation rate of women-led papers (defined here as those with women as first- and/or last- author) will be lower than expected given papers' relevant characteristics; (2) The undercitation of women-led papers will occur to a greater extent within men-led reference lists; (3) Undercitation of women-led papers will be decreasing over time, but at a slower rate within men-led reference lists; (4) Differences in undercitation between men-led and women-led reference lists will be partly explained by the structure of authors' social networks.

Results

Data description. Using Web of Science, we extracted data on research articles, reviews, and proceedings published in five top neuroscience journals since 1995. We selected the journals *Nature Neuroscience*, *Neuron*, *Brain*, *Journal of Neuroscience*, and *NeuroImage*, as they were reported by the Web of Science to have the highest Eigenfactor scores (30) among journals in the neuroscience category. In all, 61,416 articles were included in the final dataset of citing/cited papers. Full author names were provided by Web of Science for all articles published after 2006. For all articles published in 2006 or earlier, full names were drawn, when available, from Crossref or the journals' websites. To minimize missing data, we developed an algorithm to match authors for whom only first/middle initials were available to other authors in the dataset with the same initials and last name (see Methods).

Gender was assigned to first names using the 'gender' package in R (31) with the Social Security Administration (SSA) baby name dataset. For names that were not included in the SSA dataset, gender was assigned using Gender API (gender-api.com), a paid service that supports roughly 800,000 unique first names across 177 countries. We assigned 'man'('woman') to each author if their name had a probability greater than or equal to 0.70 of belonging to someone labeled as 'man'('woman') according to a given source (25). In the SSA dataset, man/woman labels correspond to the sex assigned to children at birth; in the Gender API dataset, man/woman labels correspond to a combination of sex assigned to children at birth and genders detected in social media profiles. In a random sample of 200 authors, the accuracy of these automated assignments was 0.96 (see Supplementary Information and Tables S1-S2 for further details). Gender could be assigned to both the first and last author of 88% of the papers in the dataset. Of the 12% of papers with missing data, 7% were missing because either the first- or last-author's name had uncertain gender, and 5% were missing because either the first- or last-author's name was not available. We performed the following analyses using the articles for which gender could be assigned with high probability to both authors (n = 54,226), but sensitivity analyses conducted on the full data can be found in the Supplementary Information (see Table S3).

In gender theory, sex often refers to physical attributes, as determined anatomically and physiologically, while gender often refers to a self-identity, as expressed behaviorally and in sociocultural context (32). In our analysis, the term "gender" does not refer directly to the sex of the author, as assigned at birth or chosen later, nor does it refer directly to the gender of the author, as socially assigned or self-chosen. The term "gender," in our analysis, is a function of the probability of assigned gendered names. By "woman," we mean an author whose name has a probability greater than or equal to 0.70 of being given to a child assigned female at birth or belonging to someone identifying as a woman on social media; likewise, by "man," we mean an author whose name has a probability greater than or equal to 0.70 of being given to a child assigned male at birth or belonging to someone identifying as a man on social media. The author's actual sex or gender is not identified.

Given the limitations of both probabilistic analyses and of birth assignments, the authors may in fact have a sex or gender different from the one we have assigned and/or be intersex, transgender, or nonbinary (33, 34). In some cases, citers will know the sex and/or gender of the authors they cite. In many cases, they will not know but rather infer, often via a name, the gender of the authors they cite. Instances of both known and inferred gender have the potential to incite either explicit or implicit bias in citing authors (i.e., where explicit bias involves conscious cognitive processing, implicit bias is automatic cognitive processing that presupposes social prejudices and stereotypes; 35–37). Our probabilistic analysis by gendered name therefore functions to nontrivially capture bias arising due to both known and inferred gender in citation practices.

Trends in authorship. Across the articles in the sample, the proportion of articles with a woman as first or last author significantly increased between 1995 and 2018, at a rate of roughly 0.60% per year (95% CI = [0.53, 0.67]). This trend varied across journals, with the *Journal of Neuroscience* (0.67; 95% CI = [0.57, 0.77]), *NeuroImage* (0.89; 95% CI = [0.72, 1.06]), and *Brain* (1.16; 95% CI = [0.92, 1.41]) all showing increases between 0.65% and 1.2% per

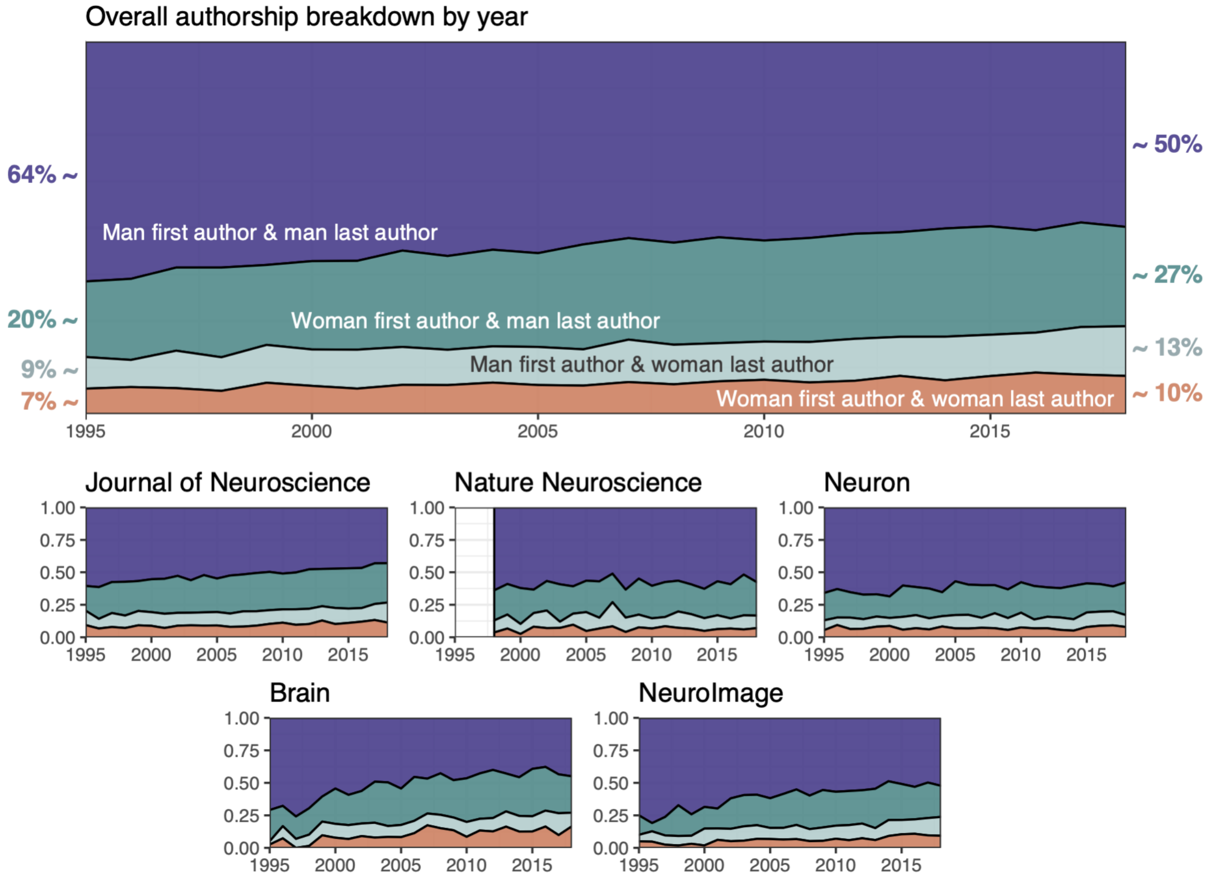

Fig. 1. Trends in author gender within top neuroscience journals between 1995 and 2018. Top panel shows the overall trends across the five journals studied. Bottom panels show the trends within each journal. From top to bottom, panels show the proportion of articles with men as first and last author (purple), women as first author and men as last author (green), men as first author and women as last author (gray), and women as both first and last author (orange). Note: *Nature Neuroscience* was not established until 1998.

year. *Neuron* showed a modest increase of 0.29% per year (95% CI = [0.12, 0.45]), and *Nature Neuroscience* did not show a clear increasing trend (0.19; 95% CI = [-0.09, 0.46]). Across these five journals, the overall proportion of articles that were either first- or last-authored by women increased from 36% in 1995 to 50% in 2018 (Figure 1).

Citation imbalance relative to overall authorship proportions. To quantify citation behavior within neuroscience articles, we specifically examined the reference lists of papers published within the past 10 years, between 2009 and 2018 (n = 28,505). Thus, while all papers in the dataset were potential *cited* papers, references to *citing* papers refer only to those published since 2009. For each citing paper, we took the subset of its citations that had been published in one of the above five journals since 1995 and determined the gender of the cited first and last authors. We removed self-citations (defined as cited papers for which either the first or last author of the citing paper was a co-author) from consideration, but see Table S4 for a comparison of self-citation rates by author gender. We then calculated the number of cited papers that fell into each of the four first author & last author categories: man & man (MM), woman & man (WM), man & woman (MW), and woman & woman (WW). As a simple first step, we compared the observed number of citations within each category to the number that would be expected if references were drawn randomly from the pool of papers (Figure 2A). To obtain the number that would be expected under this assumption of random draws, we calculated the gender proportions among all papers published prior to the citing paper – thus representing the proportion among the pool of papers that the authors could have cited – and multiplied them by the number of papers cited. The following section expands this naïve measure to account for potential relationships between author gender and other relevant characteristics of cited papers.

Of the 294,392 citations given between 2009 and 2018, MM papers received 61.8%, compared to 23.5% for WM papers, 9.0% for MW papers, and 5.7% for WW papers. The expected proportions based on the pool of citable papers were 55.4% for MM, 26.1% for WM, 10.2% for MW, and 8.2% for WW. We defined a measure of over/undercitation as the (observed % - expected %)/expected % (see Methods for further details). This measure thus represents the percent over/undercitation relative to the expected proportion. By this measure, MM papers were cited 11.6% more than expected (95% CI = [11.2, 12.0]), WM papers were cited 10.3% less than expected (95% CI = [-10.9, -9.6]), MW papers were cited 12.4% less than expected (95% CI = [-13.4, -11.2]), and WW papers were cited 30.3% less than expected (95% CI =

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.03.894378; this version posted January 11, 2020. The copyright holder for this preprint

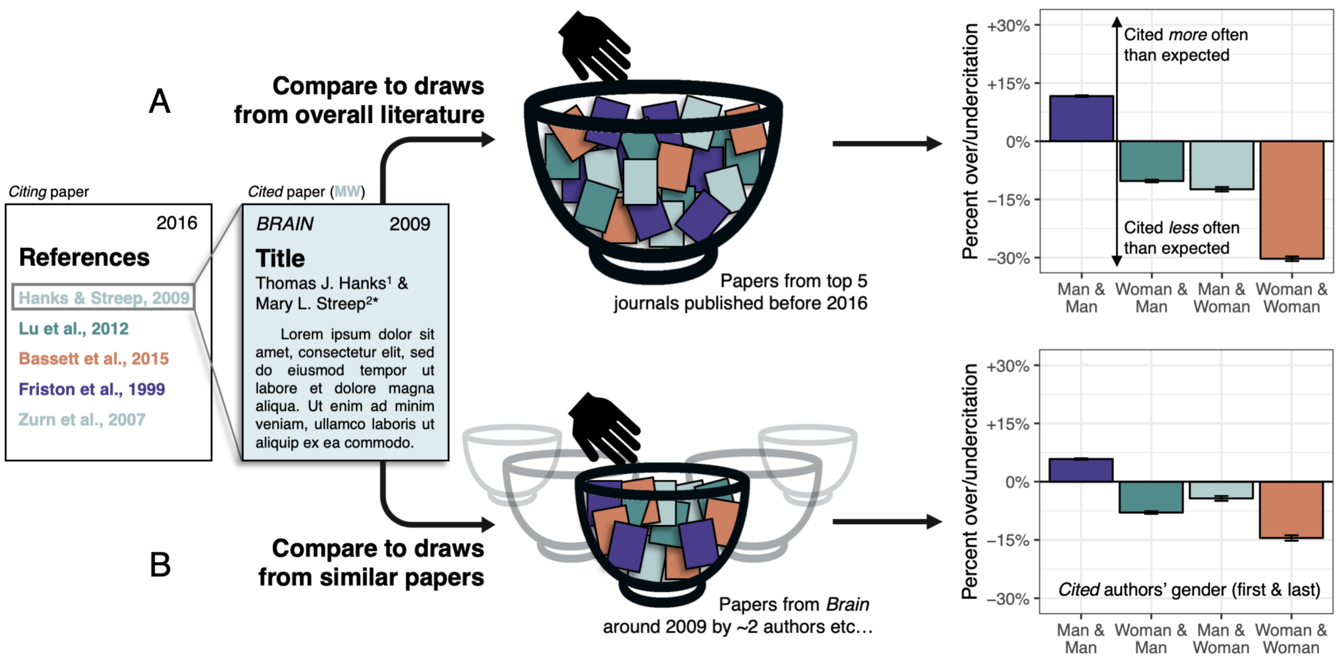

Fig. 2. Construction and visualization of over/undercitation of papers based on author gender. *(A)* Illustration of the random draws model, in which gender proportions in reference lists are compared to the overall gender proportions of the existing literature. Right panel shows the over/undercitation of different author gender groups compared to their expected proportions under the random draws model. *(B)* Illustration of the relevant characteristics model, in which gender proportions in reference lists are compared to gender proportions of articles that are similar to those that were cited across various domains. Right panel shows the over/undercitation of different author gender groups compared to their expected proportions under the relevant characteristics model.

[-31.4, -29.1]). This set of percentages correspond to MM papers being cited roughly 19,000 more times than expected, WM papers being cited roughly 7,900 fewer times than expected, MW papers being cited roughly 3,700 fewer times than expected, and WW papers being cited roughly 7,400 fewer times than expected.

Citation imbalance after accounting for papers' relevant characteristics. The comparison of citations to overall authorship proportions does not take into account other important properties of published papers that may make them more or less likely to be cited by later scholarship. The potential relationship between author gender and papers' other relevant characteristics makes it difficult to isolate the effects of gender on the rates at which work is cited. To address this issue, we developed a method for calculating the expected probabilities that a given citation would be for a MM, WM, MW, WW paper conditional on various salient characteristics of the cited paper. The characteristics of a paper that we selected as being potentially relevant for citation rates were 1) the year of publication, 2) the journal in which it was published, 3) the number of authors, 4) whether the paper was a review article, 5) the seniority of the paper's first and last authors. We then sought to compare the true citation rates to the rates that would be expected if only these non-gender characteristics were relevant.

We obtained the expected gender probabilities by specifying a generalized additive model (GAM) on the multinomial outcome of paper authorship in the four specified categories of first and last author gender. Within the GAM framework, papers' membership among these four categories was regressed on the characteristics described above (i.e., publication date, journal, author count, binary review article status, and first-/last-author seniority; see Methods for further details). Since seniority is a somewhat ambiguous concept, and is not defined in the available data, we defined authors' seniority as the number of papers on which they had been a first or last author in the time span of the study (1995-2018). Thus, the expected membership obtained for a specific article – given by the model as a set of four probabilities that sum to 1 – approximately represents the proportion of similar papers (i.e., same journal, published around the same time, etc.) that fall within each of the four gender categories.

To then estimate the gendered citation behavior of recent articles, accounting for the other relevant characteristics of cited papers, we compared the authorship gender category of each cited paper to its probabilities of belonging to each of the four categories. As opposed to the previous section, in which expected gender probabilities modeled citation as a random draw from the existing literature, the current expected probabilities model citation as a random draw from a narrow pool of papers highly similar to the cited paper (see Figure 2B).

Summing up the number of cited papers from each category again gives us the observed citation rates, and summing up the authorship gender probabilities across the cited papers gives us the new expected citation rates. As reported above, MM papers received 61.8% of citations, compared to 23.5% for WM papers, 9.0% for MW papers, and 5.7% for WW papers. Based on the relevant properties of cited papers, the expected proportions were 58.4% for MM, 25.5% for WM, 9.4% for MW, and 6.7% for WW. Thus, after accounting for

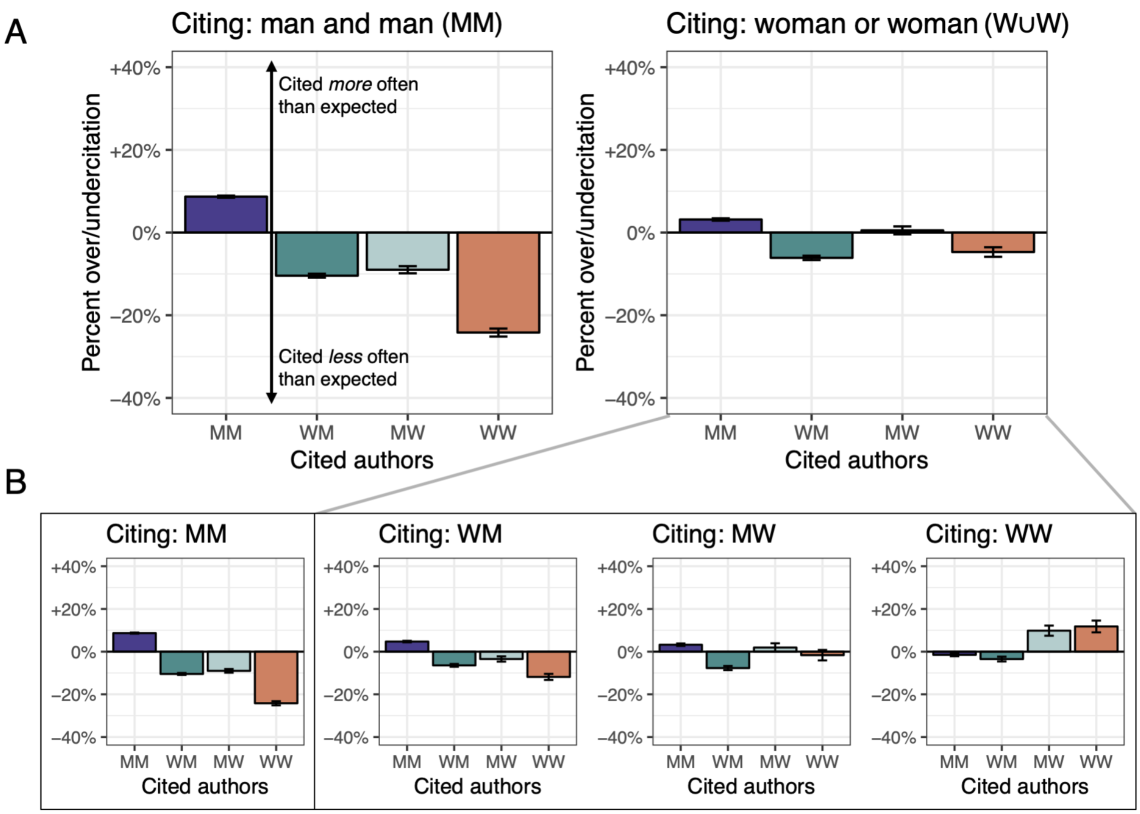

Fig. 3. Relationship between author gender and gendered citation practices. *(A)* Degree of over/undercitation of different author genders within MM reference lists (left) and within W∪W reference lists (right). Shows that papers with men as both first and last author overcite men to a greater extent than papers with women as either first or last author. *(B)* Full breakdown of gendered citation behavior within MM, WM, MW, and WW reference lists.

salient non-gender characteristics, MM papers were still cited 5.8% more than expected (95% CI = [5.5, 6.2]), WM papers were cited 8.0% less than expected (95% CI = [-8.6, -7.3]), MW papers were cited 4.3% less than expected (95% CI = [-5.4, -3.0]), and WW papers were cited 14.5% less than expected (95% CI = [-15.8, -13.2]). Of 294,392 total citations, these proportions correspond to citations being given to MM papers roughly 10,000 more times than expected, WM papers roughly 6,000 fewer times than expected, MW papers roughly 1,200 fewer times than expected, and WW papers roughly 2,900 fewer times than expected.

The effect of authors' gender on citation behavior. By focusing the present analyses on the gender make-up of reference lists, as opposed to the number of citations that articles receive, we are able to investigate the gender of the citing authors in addition to that of the cited authors. Thus, in this section we compare the gender make-up of references within papers that had men as both first and last author (referred to, as usual, as MM) to those within papers that had women as either first or last author (henceforth referred to as W∪W, comprising WM, MW, and WW papers). Of the 27,540 articles published in one of the five journals between 2009 and 2018, roughly 51% were MM and 49% were W∪W.

After separating citing articles by author gender, we find that the imbalance within reference lists shown previously is driven largely by the citation practices of MM teams. Specifically, within MM reference lists, other MM papers were cited 8.7% more than expected (95% CI = [8.2, 9.2]), WM papers were cited 10.4% less than expected (95% CI = [-11.4, -9.4]), MW papers were cited 8.9% less than expected (95% CI = [-10.7, -7.3]), and WW papers were cited 24.2% less than expected (95% CI = [-25.9, -22.3]; Figure 3A, left). Within W∪W reference lists, MM papers were cited only 3.2% more than expected (95% CI = [2.6, 3.7]), WM papers were cited 6.1% less than expected (95% CI = [-7.1, -5.1]), MW papers were cited 0.5% more than expected (95% CI = [-1.3, 2.4]), and WW papers were cited 4.7% less than expected (95% CI = [-6.9, -2.5]; Figure 3A, right).

Within the W∪W group, the citation proportions of the WM, MW, and WW subgroups suggest a more fine-grained link between the increased citation of women-led work and the increased leadership role of women on the citing team (Figure 3B). Specifically, WM teams still slightly undercite WW papers relative to expectation, but do so at roughly half the rate of MM teams. MW reference lists contain roughly the expected citation proportions across gender groups, and WW reference lists contain slightly more MW and WW papers than expected (overciting WW papers at roughly half the rate that MM teams undercite WW papers). This moderate overcitation of women-led work within women-led reference lists points to a potential role of social networks in forming authors' mental representations of the available citable work. This possibility is explored in detail in a later section.

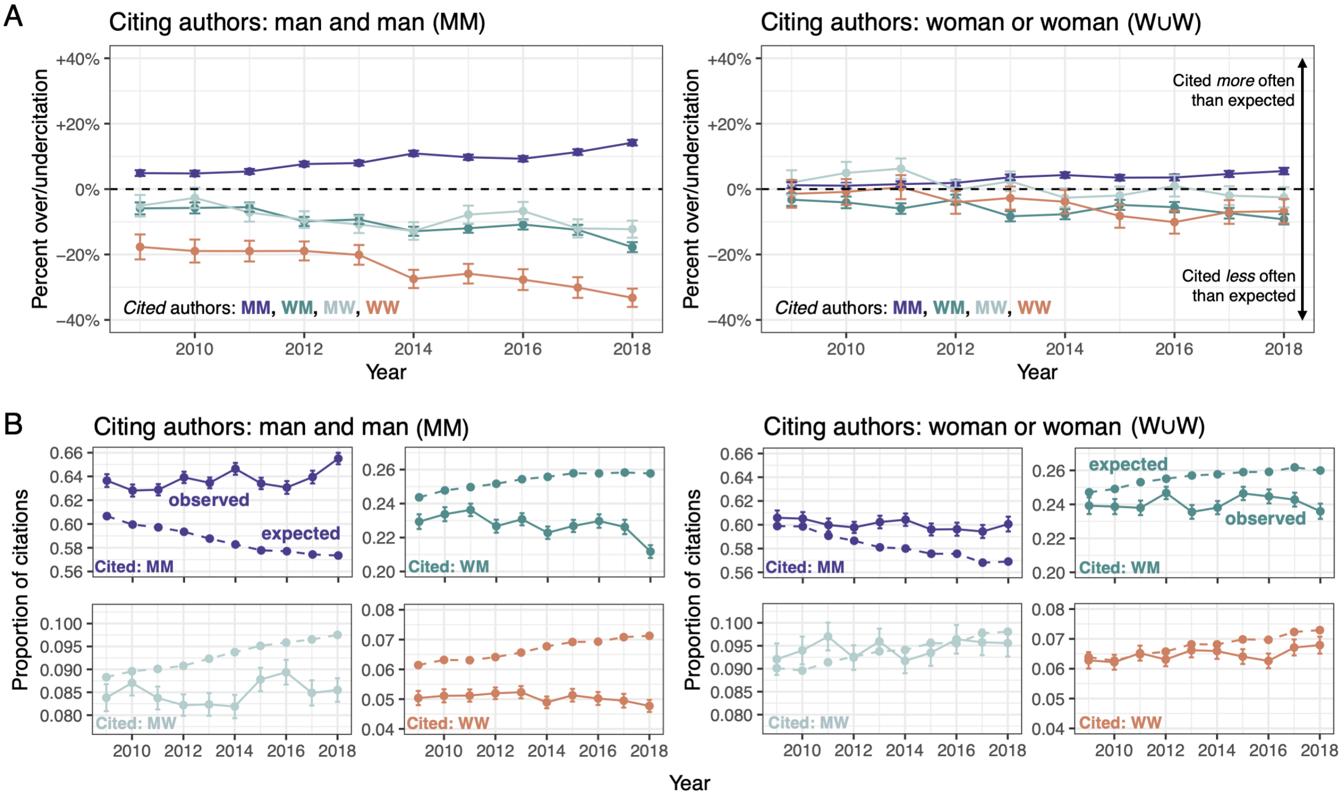

Fig. 4. Temporal trends in citation rates across cited and citing author gender. *(A)* Extent of over/undercitation across author gender categories over time, within MM (left) and W∪W (right) reference lists. *(B)* Observed (solid line) and expected (dotted line) citation proportions within MM reference lists (left) and W∪W reference lists (right). Within each section, top left shows observed/expected proportion of citations given to MM papers over time, top right shows WM papers, bottom left shows MW papers, and bottom right shows WW papers. Figure demonstrates relatively static observed proportions across groups, while expected proportions change to reflect increasing diversity within the field.

Temporal trends of citation imbalance. In addition to the overall citation behavior over the past 10 years, it was of interest to quantify the time-varying gender imbalance as the field has become more diverse over the years. As an intuitive measure of the overcitation of men in the literature, we specifically examined the absolute difference between the observed proportion of MM citations and the expected proportion of MM citations. We found that the gap between observed and expected proportions has been growing at a rate of roughly 0.40 percentage points per year (95% CI = [0.33, 0.47]). This finding suggests that citation practices are becoming less balanced over time, despite the increasing diversity of researchers.

Importantly, this growing gender gap does not simply reflect authors' propensity to cite older literature from when the field was more men-dominated, as the expected proportions account for the publication year of the articles being cited. In other words, under this construction of expected proportions, both observed and expected citation rates for each gender category would remain constant if authors were to simply cite the same literature year after year. Thus, an expanding gap between observed and expected rates suggests that either observed proportions of MM citations are increasing (e.g., authors are citing more men than they used to), expected proportions of MM citations are decreasing (e.g., authors are citing newer literature from a more diverse field), or some combination of both.

Similar to the overall tendency of MM teams to overcite other MM papers to a greater extent than W∪W teams do, we found that the degree of overcitation has been increasing faster within MM reference lists than within W∪W reference lists. Specifically, the absolute difference between the observed and expected proportions of MM citations is growing at a rate of 0.53 percentage points per year (95% CI = [0.43, 0.63]; Figure 4A, left) within MM reference lists, and it is growing at a rate of 0.27 percentage points per year (95% CI = [0.15, 0.39]; Figure 4A, right) within W∪W reference lists (difference = 0.26; 95% CI = [0.12, 0.42]). Estimates of the temporal trends within the three W∪W subgroups can be found in Table S5.

Further analysis revealed that the increasing overcitation of men in MM reference lists, and the moderately increasing overcitation of men in W∪W reference lists, reflect relatively stable citation proportions for MM papers in the face of decreasing expected proportions over time (Figure 4B). In fact, despite the increasing diversity of the field over time, the proportion of MM papers within MM reference lists has been increasing slightly, at a rate of roughly 0.15 percentage points per year (95% CI = [0.04, 0.25]). This proportion has not been clearly increasing or decreasing within W∪W reference lists, changing with a rate of -0.09 percentage points per year (95% CI = [-0.20, 0.03]).

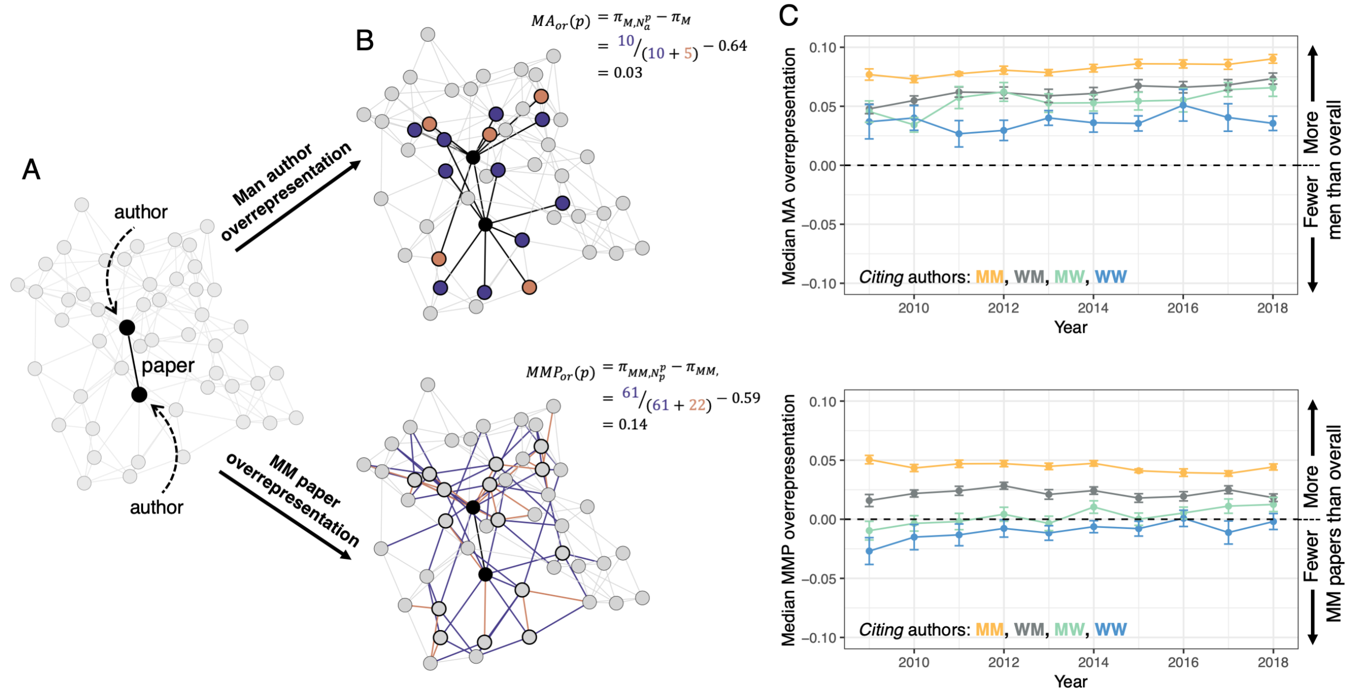

Fig. 5. Visualization of co-authorship network composition measures. *(A)* Example region of a co-authorship network, where a specific article (edge) and the first and last author (nodes) are highlighted. *(B)* Examples of the calculation of *man author overrepresentation* (MA_{or}; top) and *MM paper overrepresentation* (MMP_{or}; bottom) for the highlighted article. Here, MA_{or} is the difference between the local proportion of men (purple nodes) and the overall proportion of men. MMP_{or} is the difference between the local proportion of MM papers (purple edges) and the overall proportion of MM papers. *(C)* Differences in the local network composition based on author gender. The panel shows that papers with more women tend to have less overrepresentation of men and man-led papers within their local networks.

The relationship between social networks and citation behavior. Recent work has shown that researchers are more likely to work with other researchers of their own gender (i.e., homophily exists within co-authorship networks), and that such homophily in social networks can produce biased perceptions of the overall gender make-up of a network (38, 39). Since homophily-driven perception biases in the overall gender make-up of the field could be a potential driver of the overcitation of men by men, and slight overcitation of women by women, we sought to estimate and isolate the relationships between authors' social networks and their citation behavior. Because citations occur at the level of individual published papers, we developed two metrics to quantify gender imbalance within social networks at the paper level. Specifically, these measures consider the co-authorship network of the first and last authors of a given paper at the time of the paper's publication. Thus, two papers written by the same authors may have different values for these measures, since the co-authorship network surrounding the authors may have changed over time.

For a given paper, p, the first metric, which we refer to as man author overrepresentation, is defined as the difference between 1) the proportion of men within p's author-neighborhood (defined as the union of researchers who had previously co-authored a paper with either the first or last author of p), and 2) the overall proportion of men within the network at the time of p's publication. The second metric, which we refer to as MM paper overrepresentation, gives the difference between 1) the proportion of MM papers within p's paper-neighborhood (defined as the union of all papers written by p's first author, last author, or any of their previous co-authors), and 2) the overall proportion of MM papers within the network at the time of p's publication. Visual examples of these two measures can be seen in Figure 5A-B (see Methods for further details).

We found that across groups, co-authorship networks tended to have more men than the field as a whole, but this overrepresentation of men within co-authorship networks was especially pronounced in the networks of MM teams. Specifically, the median MM team had roughly 8.2% more men in their co-authorship network than the field's base rate (95% CI = [7.9, 8.4]), compared to the median WW team, which had roughly 3.6% more men in their network than the field's base rate (95% CI = [3.2, 4.2]). Mixed gender teams fell in the middle, with their networks being comprised of around 6% more men than the field's base rate (WM = 6.3, 95% CI = [5.9, 6.5]; MW = 5.6, 95% CI = [5.1, 6.0]; Figure 5C). The overrepresentation of MM papers among those written by authors' previous collaborators also differed based on citing authors' gender. Yet in this case, MM papers were overrepresented relative to their overall proportion only within the social networks of MM teams (+4.4%, 95% CI = [4.3, 4.5]) and WM teams (+2.2%, 95% CI = [2.0, 2.4]). MM papers were roughly proportionally represented within networks of MW teams (+0.4%, 95% CI = [0.2, 0.7]) and were slightly underrepresented within networks of WW teams (-0.8%, 95% CI = [-1.2, -0.5]; Figure 5C).

Because gendered differences in social networks tended to follow similar patterns to gendered differences in citation behavior, it was of interest to determine the degree to which

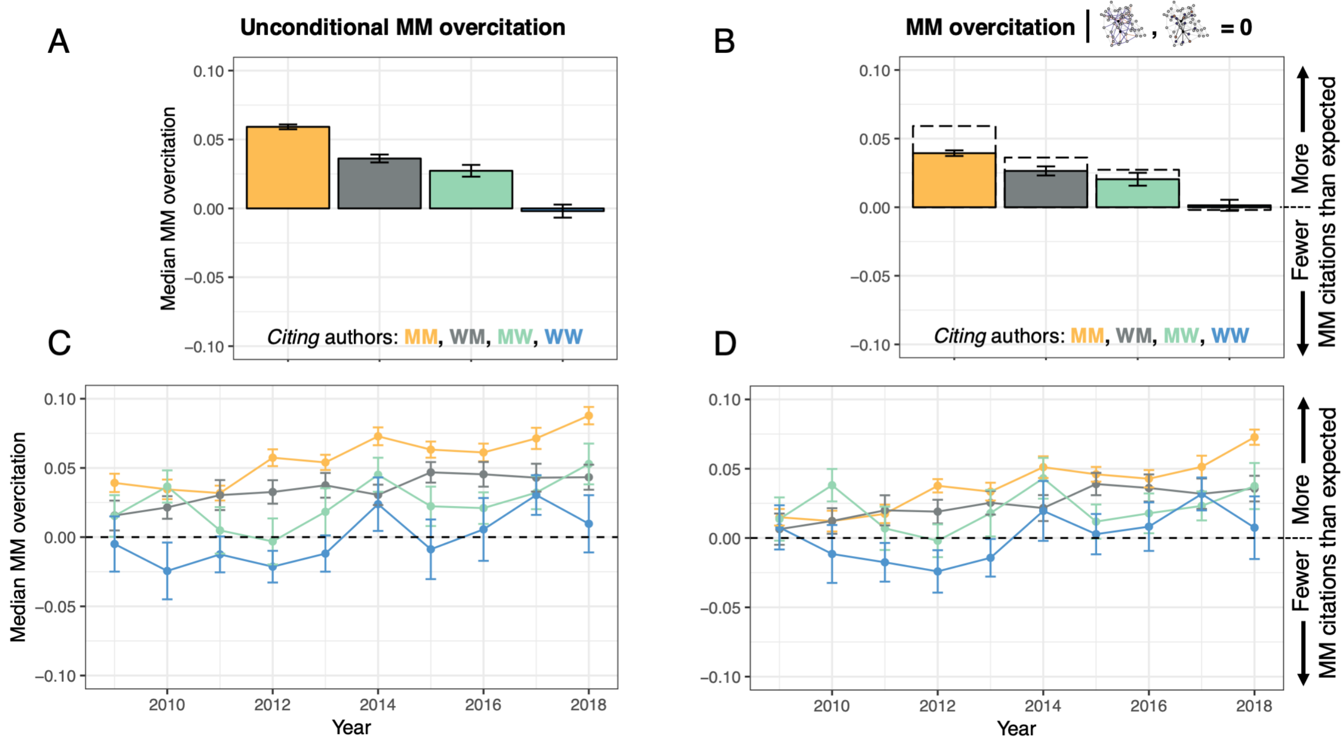

Fig. 6. Overcitation of MM papers before and after accounting for local network composition. *(A)* Overcitation of MM papers by author gender. The panel shows that MM, WM, and MW papers tend to overcite MM papers relative to expectation, while WW papers cite MM and W∪W papers at roughly the expected rate. *(B)* Overcitation of MM papers by author gender, after accounting for network effects. The panel shows that local network composition explains some of the group differences, but the general pattern remains. *(C)* Overcitation of MM papers is increasing over time across groups. *(D)* Overcitation is increasing over time across groups even after accounting for authors' social network effects.

the composition of authors' networks account for overcitation of men. For this analysis, we again utilized the absolute difference between the observed proportion of MM citations within a paper's reference list and the expected proportion based on the characteristics of the cited papers. Without accounting for differences in authors' social networks, we found that the median MM team overcites MM papers by roughly 5.9 percentage points (95% CI = [5.5, 6.3]), compared to 3.6 for WM teams (95% CI = [3.0, 4.1]), 2.7 for MW teams (95% CI = [1.9, 3.7]), and -0.2 for WW teams (95% CI = [-1.3, 0.7]; Figure 6A).

To estimate and account for the role of authors' social networks, we modeled papers' degree of MM overcitation as a function of author gender category, man author overrepresentation, and MM paper overrepresentation. Because the overcitation measure is bounded and skewed, we performed quantile regression to obtain estimates of the conditional median (see Methods for further details). Both MM paper overrepresentation and man author overrepresentation were independently associated with MM overcitation; a one percentage point increase in local overrepresentation of MM papers corresponded to a 0.24 percentage point increase in median MM overcitation (95% CI = [0.21, 0.28]), and a one percentage point increase in local overrepresentation of man authors corresponded to a 0.08 percentage point increase in median MM overcitation (95% CI = [0.04, 0.11]). Thus, the data do support the idea that there is a relationship between local coauthorship networks and citation behavior.

However, after accounting for the degree of overrepresentation of both men and MM papers within authors' social networks, differences in citation behavior remained across citing authors' gender. Specifically, conditional on authors' networks being representative of the field as a whole (i.e., local overrepresentation of men = 0 and local overrepresentation of MM papers = 0), the median MM team would still be expected to overcite MM papers by roughly 4.0 percentage points (95% CI = [3.4, 4.4]), compared to 2.7 for WM teams (95% CI = [1.8, 3.3]), 2.0 for MW teams (95% CI = [1.1, 3.0]), and 0.1 for WW teams (95% CI = [-0.8, 1.1]; Figure 6B). These results suggest that local homophily explains only part of the overcitation of men by other men. The findings also demonstrate that both before and after accounting for network effects, only WW papers tend to cite the expected proportion of MM papers, while mixed or two-man teams tend to overcite MM papers in both cases.

Notably, accounting for network effects has almost no impact on the temporal trend of MM overcitation. Specifically, we find that the degree to which MM papers are overcited in reference lists has been increasing at an identical rate of 0.44 percentage points per year both before (95% CI = [0.33, 0.55]; Figure 6C) and after (95% CI = [0.34, 0.55]; Figure 6D) accounting for network measures. This trend suggests that although social networks are associated with the magnitude of MM overcitation, they are likely not a driver of reference lists being *increasingly* unrepresentative over time.

Discussion

Like many scientific disciplines, the field of neuroscience currently faces many structural and social inequities, including marked gender imbalances (22). While the task of addressing these imbalances often depends in part on people in positions of power (e.g., journal editors (21), grant reviewers and agencies (3–5), department chairs (10–12), and presidents of scientific societies (20)), many imbalances are caused and perpetuated by researchers at all levels. One example is imbalance within citation practices (15, 16). Although the usefulness of citations as a measure of scientific value is tenuous (40), the engagement that they represent can affect how central to a field scholars are viewed to be by their peers (15). This effect on perception can then have downstream effects on conference invitations, grant and fellowship awards, tenure and promotion, inclusion in syllabi, and even student evaluations. As a result, understanding and eliminating gender bias in citation practices is vital for addressing gender imbalances in a field.

In this study, we sought to determine whether there is evidence of gender bias in neuroscience citations, and whether that bias itself differs based on the gender of the citing authors. We indeed found evidence that neuroscience reference lists tend to include more papers with men as first and last author than would be expected if gender was not a factor. Importantly, this overcitation of men and undercitation of women is driven largely by the citation practices of men. Specifically, papers with men as first and last author overcite other man/man papers by 9% relative to the expected proportion, undercite woman/man papers by 10%, undercite man/woman papers by 9%, and undercite woman/woman papers by 24%. Papers with women in one or both primary authorship positions overcite man/man papers by 3%, undercite woman/man papers by 6%, overcite man/woman papers by 0.5%, and undercite woman/woman papers by 5%. These results are consistent with results from other fields that show that men are less likely to cite work by women (15, 16, 25).

Gender inequity in general — and, one might therefore argue, gender inequity in citational practices — is understood to result from both systemic bias and individual bias. Systemic bias, also known as institutional bias, refers to discriminatory values, practices, and mechanisms that function at the intergroup level in the domain of social institutions (41). Individual bias may be either explicit or implicit. Explicit bias is consciously held or expressed prejudice against a particular group or an individual of that group, resulting in material, psychological, or physical harms (42). Implicit bias, on the other hand, is a set of subconsciously harbored discriminatory attitudes against a particular group or an individual of that group, which can result in prejudicial speech and social behaviors (37, 43). Implicit bias is traceable in individual attitudes (and resulting actions) relative to something as concrete as physical appearance and as abstract as a mere name. Indeed, implicit bias with respect to names has been shown in studies of race-based (44, 45) and gender (7, 46) discrimination. The undercitation of women in neuroscience papers, therefore, may be due to systemic gender bias or to explicit or implicit individual bias relative either to the known gender of an author or to an author's gendered name. Our analysis thus extends and contributes to existing literature on bias, gender inequity, and citational practices.

Recent work has also shown that homophily in social networks (i.e., an increased likelihood of being connected to people of the same gender) can lead to biases in individuals' perceptions of the overall proportion of men and women (39). In this context, such an effect could imply that men's preferential attachment to other men could lead them to overestimate the base rates of men within the field. To determine the degree to which homophily in researchers' social networks explains the prevalence of men overciting other men, we developed a co-authorship network and measured the degree to which men (and man-led papers) were overrepresented in authors' local neighborhoods. We find that while these features are indeed associated with authors' overcitation of men, man/man papers are still overcited in the reference lists of other man/man papers, and to a lesser extent in reference lists of woman/man and man/woman papers, even after accounting for network effects. The reference lists of woman/woman papers are the only group that tend to cite man/man papers at the expected rate. These results, while consistent with the presence of some homophily effect, show that men tend to overcite other men even if their social networks are representative of the field.

There are several possible mechanisms that could explain the remaining difference in citation behavior between man-led teams and woman-led teams. One obvious explanation is greater conscious or unconscious bias among men, which could lead them to evaluate woman-led work more harshly and thus be more hesitant to cite such work. This explanation would be consistent with studies that have shown evaluative bias in the realms of graduate admissions (47), faculty hiring (2), grant funding (5), and promotion (12).

Other explanations, like the overrepresentation of men in course syllabi (48) and in conference speaking roles (22), could partly explain the difference between groups (e.g., women may take more courses taught by women, who discuss and assign more work by women). Yet mechanisms like these would likely be more consistent with an overall overcitation of men that does not differ based on the gender of the citing authors. In that case, the fact that teams with more women display less gender citation bias could be explained by their conscious efforts to seek out and cite work by other women. If this is the case, it is plausible that women's attempts to address gender imbalance could make their citation practices more representative of their fields, while men's indifference or lack of awareness could lead to the propagation of imbalances present in syllabi and conferences.

Regardless of the mechanisms that drive these imbalances, greater awareness of existing (and persisting, and even increasing) imbalances in citation practices is likely an important step in heightening researchers' willingness to address these issues. Recent work has laid out guidelines for responsible citation practices (49), which include consideration of gender imbalance. There also exist tools to probabilistically measure the proportion of women and men within course syllabi and reference lists (50). Various organizations also provide information that can assist researchers in creating representative reference lists. These include BiasWatchNeuro (biaswatchneuro.com), which publishes base rates for different subfields within neuroscience, and Women in Neuroscience (winrepo.org) and Anne's List (anneslist.net), which contain detailed, searchable databases of women in neuroscience and their areas of expertise. Neuroscience might also consider having a reference list that represents additional axes of marginalization (e.g. gender, race, class, sexuality, disability, citizenship, etc.), given both the intersectional discrimination of women in the academy (51) and the aspiration to address social inequities in the field more broadly. The American Philosophical Association's UP Directory (52) provides a potential reference point for this type of inclusive list. Addressing the identified imbalances will require researchers, particularly men, to make use of such resources and engage in more thoughtful citation practices. Educating graduate students about citation practices will also be vital, and such discussions could potentially be incorporated directly into the NIH and NSF's "responsible conduct of research (RCR)" requirements.

Beyond a growing individual and collective thoughtfulness, paired with ad-hoc efforts, to achieve gender balance in neuroscience reference lists, the ethics of citation practices remain to be further defined. Righting social inequities may be accomplished on a number of different models. On the distributive model, for example, justice refers to the morally proper distribution of social goods and resources or, in this case, citations (53). Exactly how that distribution ought to be circumscribed, however, remains in question. On the equality-based distributive model, citations ought to be allocated to all authors equally; while on the equity-based distributive model, citations ought to be allocated to authors differentially based upon select factors, which may include merit, need, or authority (54, 55). The distributive paradigm on the whole, however, is limited insofar as it emphasizes commodity parity across economies of exchange over differential responsibility for histories and structures of inequality (56, 57). Difference models, by contrast, recommend acts of reparative justice (58), which might include affirmative action (59) in citational practices, institutional reform to support citation parity, and disciplinary redress of gender bias more broadly.

Distributive and difference-based models raise a series of important questions for citation ethics in neuroscience. Should gender balance in citation practices reflect random distributions or distributions tuned to relevant features (and, if so, which features)? Are such distributional structures sufficient either to correct for a history of underrepresentation or to secure a future of equitable representation? Given the lassitude with which social change occurs—and the worsening of gender imbalance in citation practices in neuroscience overall—is it justifiable for some research teams to significantly overcite papers produced by women-led teams wherever possible? Might the effects of systemic gender bias on undercitation practices be counteracted by the significant employment of women in field-specific decision-making bodies, reforming checkpoints, and professional activism? And, given the function of implicit bias and its capacity for correction via experience, should researchers of all genders commit to collaborating more robustly with women and other gender minorities?

Overall, the work of citation is an important element in the research ethics of any field. Insofar as citation patterns today have inescapable effects on the future of neuroscience, citational practices in the field warrant more serious attention.

Limitations and future work. This work is subject to several important limitations. First, this study focuses on citing and cited articles published in five top neuroscience journals. Although this focus has the benefit of reducing the confounding effects of journal prestige, it also may lead to artificially low proportions of women-led work if there are biases in prestige publishing. Future work would be needed to understand the role of more specialized journals. Second, this study does not address the potential effects of authors' institutional prestige. As a result, a combination of gender biases in hiring and prestige-based citation behavior could introduce confounds. Future work could attempt to quantify and isolate the effect of departmental prestige in this context. Third, the methods used for gender determination are limited to binary man/woman gender assignments. This study design, therefore, is not well accommodated to intersex, transgender, and/or non-binary identities, and incorrectly assumes that all authors in the dataset can be placed into one of two categories. Ideally, future work will be able to move beyond the gender binary, potentially by applying methods that utilize pronouns or other forms of self-identification. Finally, the current study investigates biases solely along gender lines. Future work could extend these types of analyses to examine biases along, for example, race or ethnicity, as well as their intersection with gender.

Methods

Data collection. We drew data for this study from the Thomson Reuters' Web of Science (WoS) database. The Web of Science database indexes neuroscience journals according to the Science Citation Index Expanded, and we selected the neuroscience journals with the five highest Eigenfactor scores for study. Eigenfactor scores give a count of incoming citations, where citations are weighted by the impact of the citing journal. Therefore, this measure roughly mimics the classic version of Google page rank, and attempts to characterize the influence a journal has within its field (30, 60). The journals selected were *Brain*, *Journal of Neuroscience*, *Nature Neuroscience*, *NeuroImage*, and *Neuron*.

All articles published between 1995 and 2018 were downloaded, and articles classified as articles, review articles, or proceedings papers that were labeled with a Digital Object Identifier (DOI) were included in the analyses. The data downloaded from WoS included papers' author names, reference lists, publication dates, and DOI, and we obtained information on each paper's referencing behavior by matching DOIs contained within a reference list to DOIs of papers included in the dataset.

Although authors' last names were included for all papers, authors' first names were only regularly included in the data for papers published after 2006. For all papers published in or before 2006, we searched for author first names using Crossref's API. When first names were not available on Crossref, we searched for them on the journals' webpage for the given article. To minimize the number of papers for which we only had access to authors' initials, to remove self-citations, and to develop a co-authorship network, we implemented a name disambiguation algorithm.

Author name disambiguation. To minimize missing data, allow for name gender assignment, and allow for author matching across papers, we implemented an algorithm to disambiguate authors for whom different versions of their given name or initials were available across papers. We began by separating first and last names according to the method used by the given source (e.g., WoS typically used "last, first; last, first"). We then identified cases in which only initials were available after the previously described searching steps by marking authors for whom the first name entry contained only uppercase letters (as we found that many initials-only entries did not contain periods).

For each case, we collected all other entries that contained the same first/middle initials and the same last name. If only one unique first/middle name matched the initials of the given entry, or if distinct matches were all variants of the same name, we assigned that name to the initials. If there were multiple names in the dataset that fit the initial/last name combination of the given entry, then we did not assign a name to the initials. For example, if an entry listed an author as R. J. Dolan, and we found matches under Ray J. Dolan and Raymond J. Dolan, we would replace the R. J. Dolan entry with the more common completed variant. If, instead, we found matches under Ray J. Dolan and Rebecca J. Dolan, we would not assign a name to the original R. J. Dolan entry.

Next we matched different name variants for the sake of tracking individual authors across their papers. To find and connect variants, we searched for instances of author entries with matching last names and either the same first name or first names that were listed as being commonly used nicknames according to the Secure Open Enterprise Master Patient Index (61). If there were no matches that fit that description, the name was retained. If there was one match that occurred more commonly, the less common variant was changed to the more common variant. If there were multiple matches that did not have any conflicting initials (some having a middle initial and others not having one was not considered conflicting), then less common variants were changed to the more common variant. If there were multiple matches that did have conflicting initials (e.g., Ray Dolan being matched to both Raymond S. Dolan and Raymond J. Dolan), then the target name was not changed.

Author gender determination. For all authors with available first names, we carried out the process of gender determination in two steps. First, we used the Social Security Administration (SSA) database as implemented in the 'gender' R package (31), which returns the proportion of matching baby names given to infants assigned female or male at birth between the years 1932 and 2012. The assignment of "woman" was given to names with greater than a 0.7 probability of being assigned female at birth, and the assignment of "man" was given to names with greater than a 0.7 probability of being assigned male at birth. Because this database is primarily useful for United States-based authors, we used Gender API as a secondary source for any names that had a probability between 0.3 and 0.7 or were not present in the SSA database. To determine the extent of potential gender mislabeling, we conducted a manual study on a sample of 200 authors. The relative accuracy of the automated determination procedure at the level of both individual authors (Accuracy ≈ 0.96; Table S1) and article gender categories (Accuracy ≈ 0.92; Table S2) can be seen in the Supplementary Information.

Statistical analysis. In this section, we describe the formal statistical analysis that we used to address four distinct hypotheses. In each subsection we state the hypothesis first, followed by the statistical analysis used to test it.

Hypothesis 1: The overall citation rate of women-led papers will be lower than expected given papers' relevant characteristics. To test this hypothesis, we first estimated the expected gender of papers' first and last authors conditional on several characteristics deemed to be potentially related to citation rates. We fit a generalized additive model (GAM) on the multinomial outcome {MM, WM, MW, WW}, in which the features of interest were 1) month and year of publication, 2) combined number of publications by the first and last authors, 3) number of total authors on the paper, 4) the journal in which it was published, and 5) whether it was a review paper. The GAM was fit using the 'mgcv' package in R (62), using penalized thin plate regression splines for estimating smooth terms of publication date, author experience, and team size.

This procedure resulted in each paper in the dataset having an expected probability of being written by a MM, WM, MW, or WW team given its various characteristics. A paper's expected probability vector can then be thought to represent the approximate gender proportions among articles that are highly similar to it across the five dimensions listed above. Using these probabilities, we were able to calculate the expected number of citations given to each group. We performed this calculation by summing over the expected probabilities for all of the papers contained within the reference lists of papers published between 2009 and 2018.

To calculate the observed number of citations given to each group, we simply summed over the {MM, WM, MW, WW} dummy variable for all of the papers contained within the reference lists of papers published between 2009 and 2018. These values were compared by calculating the percent difference from expectation for each author gender group. For example, for WW papers, this percent change in citation would be defined as,

$$\Delta_{WW} = \frac{obs_{WW} - exp_{WW}}{exp_{WW}},$$

where obs_{WW} is the number of citations given to WW papers between 2009 and 2018, and exp_{WW} is the expected number of citations given to WW papers between 2009 and 2018.

Confidence intervals for these values were calculated by bootstrapping citing papers. As opposed to bootstrapping individual instances of citations, this method maintains the dependence structure of the clusters of cited articles within citing articles. Notably, performing the summation over all citations results in the upweighting of articles with many citations, and the downweighting of articles with few. This approach helps to improve the stability of the estimates, but could potentially introduce confounds. To determine the impact of this decision, we conducted sensitivity analyses in which we used the mean of article-level over/undercitation, and found little difference between the two estimation strategies (see Supplementary Information and Table S6 for further details).

Hypothesis 2: The undercitation of women-led papers will occur to a greater extent within men-led reference lists. To test this hypothesis, we used very similar metrics to those described in the previous section. The primary difference is that instead of calculating the observed and expected citations by summing over the citations within all reference lists between 2009 and 2018, in this section we performed those summations separately for reference lists in papers with men as first and last author (MM papers) and papers with women as first or last author (W∪W papers). For example, to estimate the over/undercitation of WW papers within the reference lists of MM papers, we define,

$$\Delta_{WW}^{(MM)} = \frac{obs_{WW}^{(MM)} - exp_{WW}^{(MM)}}{exp_{WW}^{(MM)}},$$

where $obs_{WW}^{(MM)}$ is the total number of citations given to WW papers within MM reference lists, and $exp_{WW}^{(MM)}$ is the expected number of citations given to WW papers within MM reference lists.

Hypothesis 3: Undercitation of women-led papers will be decreasing over time, but at a slower rate within men-led reference lists. As there are four separate measures representing over or undercitation of each author group, we calculated change in the overcitation of men over time using the simple measure of the absolute difference between the observed proportion of MM papers cited and the expected proportion of MM papers cited. This measure of change is given by,

$$\delta_{MM,year} = \frac{obs_{MM,year} - exp_{MM,year}}{obs_{year}},$$

where obs_{year} is the total number of citations within a given year, $obs_{MM,year}$ is the number of citations given to MM papers in a specific year, and $exp_{MM,year}$ is the expected number of citations given to MM papers in a specific year. The change in the overcitation of men over time is estimated using a linear regression of $\delta_{MM,year}$ on year, and the confidence interval of this estimate is obtained using the same article bootstrap procedure previously described.

Similarly, to estimate the change in overcitation of MM papers separately within MM reference lists and W∪W reference lists, we defined group-specific measures of yearly overcitation. For example, overcitation of MM papers within MM reference lists for a specific year would be given by,

$$\delta_{MM,year}^{(MM)} = \frac{obs_{MM,year}^{(MM)} - exp_{MM,year}^{(MM)}}{obs_{year}^{(MM)}},$$

where $obs_{year}^{(MM)}$ is the total number of citations within MM reference lists in a specific year, $obs_{MM,year}^{(MM)}$ is the number of citations given to MM papers within MM reference lists in a specific year, and $exp_{MM,year}^{(MM)}$ is the expected number of citations given to MM papers within MM reference lists in a specific year.

Hypothesis 4: Differences in undercitation between men-led and women-led reference lists will be partly explained by the structure of authors' social networks. To test this hypothesis, we developed a temporal co-authorship network in which nodes were individual authors (only authors who appeared as first or last author in at least one paper in the dataset were included), and binary edges represented the fact that two authors had appeared on at least one paper together prior to a given date. It is of interest in this section to estimate the relationship between authors' local network composition and their citation behavior. Because citation behavior occurs at the level of a reference list within a specific paper with both a first and a last author (rather than at the level of a single node, or author), we sought to define two measures of local network composition at the paper level. For the purposes of these analyses, we consider a paper to be the set $\{a_f, a_l, m\}$, where a_f is the first author, a_l is the last author, and m is the month of publication. We then define a paper's local neighborhood of authors, N_a^p, to be the authors that are connected by shared publication to either a_f or a_l prior to month m. We also define a paper's local neighborhood of papers, N_p^p, to be the union of all papers authored by anyone within N_a^p prior to month m.

The two measures of local network composition are man author overrepresentation and MM paper overrepresentation. We define man author overrepresentation as the difference between the proportion of men within a paper's local author neighborhood, N_a^p, and that of the overall network. For paper p, this measure is therefore given by,

$$MA_{or}(p) = \pi_{M,N_a^p} - \pi_M,$$

where π_M is the proportion of men in the full co-authorship network, and π_{M,N_a^p} is the proportion of men within paper p's local author neighborhood. Similarly, we define MM paper overrepresentation as the difference between the proportion of MM articles within a paper's local paper neighborhood, N_p^p, and that of the overall network. For paper p, this measure is therefore given by,

$$MMP_{or}(p) = \pi_{MM,N_p^p} - \pi_{MM},$$

where π_{MM} is the overall proportion of MM articles within the data, and π_{MM,N_p^p} is the proportion of MM articles within paper p's local paper neighborhood.

To estimate the relationship between these metrics and the degree of overcitation of men within reference lists, we defined a paper-level measure of the absolute difference between the observed and expected proportion of MM papers. Similar to the previously described $\delta_{MM,year}^{(MM)}$ measure that quantified the overcitation of MM papers within all MM reference lists from a given year, here we define a measure of overcitation within an individual paper, p. It is given by,

$$\delta_{MM}^{(p)} = \frac{obs_{MM}^{(p)} - exp_{MM}^{(p)}}{obs^{(p)}},$$

where $obs_{MM}^{(p)}$ is the number of MM citations within paper p's reference list, $exp_{MM}^{(p)}$ is the expected number of MM citations within paper p's reference list based on the GAM-estimated assignment probabilities of each cited paper, and $obs^{(p)}$ is the total number of candidate citations within paper p's reference list.

The relationships between $\delta_{MM}^{(p)}$, $MMP_{or}(p)$, $MA_{or}(p)$, and $\{MM, WM, MW, WW\}$ are estimated using weighted quantile regression, with the MM overcitation metric, $\delta_{MM}^{(p)}$, as the outcome. We performed quantile regression because of the bounded and skewed nature of the $\delta_{MM}^{(p)}$ measure, but the results of a sensitivity analysis using linear regression can be found in Table S7. We define the weights to be equal to the number of candidate citations within a given paper's reference list; this choice gives higher weight to papers for which the outcome is more stable. Results from an unweighted model can be found in Table S6. We also take the τ value of the quantile regression formula to be 0.5, resulting in a model fit to the median of the outcomes. Confidence intervals are again obtained by the article bootstrap method.

Acknowledgements

We thank David Lydon-Staley and Dale Zhou for constructive comments on an earlier version of this manuscript. RTS would like to acknowledge support from the National Institute of Neurological Disorders and Stroke (R01 NS085211 & R01 NS060910). DSB would like to acknowledge support from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, and the National Science Foundation CAREER (PHY-1554488). The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Citation gender diversity statement

The gender balance of papers cited within this work was quantified using a combination of automated gender-api.com estimation and manual gender determination from authors' publicly available pronouns. Among the 60 cited works that had named authors, 30% ($n = 18$) were MM, 10% ($n = 6$) were WM, 13% ($n = 8$) were MW, and 47% ($n = 28$) were WW.

References

1. Luke Holman, Devi Stuart-Fox, and Cindy E. Hauser. The gender gap in science: How long until women are equally represented? *PLOS Biology*, 16(4):e2004956, April 2018. ISSN 1545-7885. doi: 10.1371/journal.pbio.2004956.
2. Corinne A. Moss-Racusin, John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41):16474–16479, October 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1211286109.
3. Lutz Bornmann, Rudiger Mutz, and Hans-Dieter Daniel. Gender differences in grant peer review: A meta-analysis. *Journal of Informetrics*, 1(3):226–238, July 2007. ISSN 17511577. doi: 10.1016/j.joi.2007.03.001.
4. Reshma Jagsi. Sex Differences in Attainment of Independent Funding by Career Development Awardees. *Annals of Internal Medicine*, 151(11):804, December 2009. ISSN 0003-4819. doi: 10.7326/0003-4819-151-11-200912010-00009.
5. Romy van der Lee and Naomi Ellemers. Gender contributes to personal research funding success in The Netherlands. *Proceedings of the National Academy of Sciences*, 112(40): 12349–12353, October 2015. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1510159112.
6. Heather Sarsons. Recognition for Group Work: Gender Differences in Academia. *American Economic Review*, 107(5):141–145, May 2017. ISSN 0002-8282. doi: 10.1257/aer.p20171126.
7. Lillian MacNell, Adam Driscoll, and Andrea N. Hunt. What's in a Name: Exposing Gender Bias in Student Ratings of Teaching. *Innovative Higher Education*, 40(4):291–303, August 2015. ISSN 0742-5627, 1573-1758. doi: 10.1007/s10755-014-9313-4.
8. Friederike Mengel, Jan Sauermann, and Ulf Zölitz. Gender Bias in Teaching Evaluations. *Journal of the European Economic Association*, 17(2):535–566, April 2019. ISSN 1542-4766, 1542-4774. doi: 10.1093/jeea/jvx057.
9. Anne Boring. Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145:27–41, January 2017. ISSN 00472727. doi: 10.1016/j.jpubeco.2016.11.006.
10. Mathias W. Nielsen. Limits to meritocracy? Gender in academic recruitment and promotion processes. *Science and Public Policy*, 43(3):386–399, June 2016. ISSN 0302-3427, 1471-5430. doi: 10.1093/scipol/scv052.
11. Marieke Van den Brink. Scouting for talent: Appointment practices of women professors in academic medicine. *Social Science & Medicine*, 72(12):2033–2040, June 2011. ISSN 02779536. doi: 10.1016/j.socscimed.2011.04.016.
12. Maria De Paola and Vincenzo Scoppa. Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia. *Economica*, 82(325):162–188, January 2015. ISSN 00130427. doi: 10.1111/ecca.12107.
13. Isaac Wilhelm, Sherri Lynn Conklin, and Nicole Hassoun. New data on the representation of women in philosophy journals: 2004–2015. *Philosophical Studies*, 175(6):1441–1464, June 2018. ISSN 0031-8116, 1573-0883. doi: 10.1007/s11098-017-0919-0.
14. Vincent Larivière, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. Bibliometrics: Global gender disparities in science. *Nature*, 504(7479):211–213, December 2013. ISSN 0028-0836, 1476-4687. doi: 10.1038/504211a.
15. Marianne A. Ferber and Michael Brun. The Gender Gap in Citations: Does It Persist? *Feminist Economics*, 17(1):151–158, January 2011. ISSN 1354-5701, 1466-4372. doi: 10.1080/13545701.2010.541857.
16. Daniel Maliniak, Ryan Powers, and Barbara F. Walter. The Gender Citation Gap in International Relations. *International Organization*, 67(4):889–922, October 2013. ISSN 0020-8183, 1531-5088. doi: 10.1017/S0020818313000209.
17. Neven Caplar, Sandro Tacchella, and Simon Birrer. Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy*, 1(6):0141, June 2017. ISSN 2397-3366. doi: 10.1038/s41550-017-0141.
18. Di Fang, Ernest Moy, Lois Colburn, and Jeanne Hurley. Racial and Ethnic Disparities in Faculty Promotion in Academic Medicine. *JAMA*, 284(9):1085–1092, September 2000. ISSN 0098-7484. doi: 10.1001/jama.284.9.1085.
19. Alexander M. Petersen, Santo Fortunato, Raj K. Pan, Kimmo Kaski, Orion Penner, Armando Rungi, Massimo Riccaboni, H. Eugene Stanley, and Fabio Pammolli. Reputation and impact in academic careers. *Proceedings of the National Academy of Sciences*, 111(43):15316–15321, October 2014. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1323111111.
20. Marian Joels and Carol Mason. A Tale of Two Sexes. *Neuron*, 82(6):1196–1199, June 2014. ISSN 08966273. doi: 10.1016/j.neuron.2014.05.021.
21. Promoting diversity in neuroscience. *Nature Neuroscience*, 21(1):1–1, January 2018. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-017-0052-6.
22. Jessica Schrouff, Doris Pischedda, Sarah Genon, Gregory Fryns, Ana Luísa Pinho, Eliana Vassena, Antonietta G. Liuzzi, and Fabio S. Ferreira. Gender bias in (neuro)science: Facts, consequences, and solutions. *European Journal of Neuroscience*, March 2019. ISSN 0953816X. doi: 10.1111/ejn.14397.
23. Paula Chakravartty, Rachel Kuo, Victoria Grubbs, and Charlton McIlwain. #CommunicationSoWhite. *Journal of Communication*, 68(2):254–266, April 2018. ISSN 0021-9916, 1460-2466. doi: 10.1093/joc/jqy003.
24. Yannik Thiem, Kris F. Sealey, Amy E. Ferrer, Adriel M. Trott, and Rebecca Kennison. Just Ideas? The Status and Future of Publication Ethics in Philosophy: A White Paper. Technical report, 2018.
25. Michelle L. Dion, Jane Lawrence Sumner, and Sara McLaughlin Mitchell. Gendered Citation Patterns across Political Science and Social Science Methodology Fields. *Political Analysis*, 26(3):312–327, July 2018. ISSN 1047-1987, 1476-4989. doi: 10.1017/pan.2018.12.
26. Margaret W. Rossiter. The Matthew Matilda Effect in Science. *Social Studies of Science*, 23 (2):325–341, May 1993. ISSN 0306-3127, 1460-3659. doi: 10.1177/030631293023002004.
27. Scott Selisker. The Bechdel Test and the Social Form of Character Networks. *New Literary History*, 46(3):505–523, 2015. ISSN 1080-661X. doi: 10.1353/nlh.2015.0024.
28. David Garcia, Ingmar Weber, and Venkata Garimella. Gender asymmetries in reality and fiction: The bechdel test of social media. *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, 04 2014.
29. Sara McLaughlin Mitchell, Samantha Lange, and Holly Brus. Gendered Citation Patterns in International Relations Journals1. *International Studies Perspectives*, 14(4):485–492, 2013. ISSN 1528-3577. doi: 10.1111/insp.12026.
30. Carl T. Bergstrom, Jevin D. West, and Marc A. Wiseman. The EigenfactorTM Metrics. *Journal of Neuroscience*, 28(45):11433–11434, November 2008. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0003-08.2008.
31. Cameron Blevins and Lincoln Mullen. Jane, John ... Leslie? A Historical Method for Algorithmic Gender Prediction. *Digital Humanities Quarterly*, 9(3), 2015.
32. Anne Fausto-Sterling. *Sexing the body: gender politics and the construction of sexuality*. Basic Books, New York, NY, 1st ed edition, 2000. ISBN 978-0-465-07713-7 978-0-465-07714-4. OCLC: ocm43468482.
33. Ellen K. Feder. *Making sense of intersex: changing ethical perspectives in biomedicine*. Indiana University Press, Bloomington, 2014. ISBN 978-0-253-01224-1 978-0-253-01228-9.
34. Susan Stryker. *Transgender history*. Seal studies. Seal Press : Distributed by Publishers Group West, Berkeley, CA, 2008. ISBN 978-1-58005-224-5. OCLC: ocn183914566.
35. Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz. Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480, 1998. ISSN 1939-1315, 0022-3514. doi: 10.1037/0022-3514.74.6.1464.
36. Marilynn B. Brewer. The Psychology of Prejudice: Ingroup Love and Outgroup Hate? *Journal of Social Issues*, 55(3):429–444, January 1999. ISSN 0022-4537, 1540-4560. doi: 10.1111/0022-4537.00126.
37. Michael Brownstein. Implicit bias. In *The Stanford Encyclopedia of Philsophy*. E. Zalta (Ed.), fall 2019 edition edition, 2019.
38. Luke Holman and Claire Morandin. Researchers collaborate with same-gendered colleagues more often than expected across the life sciences. *PLOS ONE*, 14(4):e0216128, April 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0216128.
39. Eun Lee, Fariba Karimi, Claudia Wagner, Hang-Hyun Jo, Markus Strohmaier, and Mirta Galesic. Homophily and minority-group size explain perception biases in social networks. *Nature Human Behaviour*, August 2019. ISSN 2397-3374. doi: 10.1038/s41562-019-0677-4.
40. Dag W. Aksnes, Liv Langfeldt, and Paul Wouters. Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. *SAGE Open*, 9 (1):215824401982957, January 2019. ISSN 2158-2440, 2158-2440. doi: 10.1177/2158244019829575.
41. PJ Henry. *Henry, P. J. (2010). Institutional bias. In J. F. Dovidio, M. Hewstone, P. Glick, V. M. Esses (Eds.), Handbook of prejudice, stereotyping, and discrimination (426-440). London: Sage.*, pages 426–440. 01 2010.
42. Jessica A. Clarke. Explicit Bias. *Northwestern University Law Review*, 113(3):505–586.
43. Anthony G. Greenwald and Mahzarin R. Banaji. Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1):4–27, 1995. ISSN 1939-1471, 0033-295X. doi: 10.1037/0033-295X.102.1.4.
44. Wendy Conaway and Sonja Bethune. Implicit Bias and First Name Stereotypes: What are the Implications for Online Instruction? *Online Learning*, 19(3), March 2015. ISSN 2472-5730, 2472-5749. doi: 10.24059/olj.v19i3.452.
45. Marianne Bertrand and Sendhil Mullainathan. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4):991–1013, August 2004. ISSN 0002-8282. doi: 10.1257/0002828042002561.
46. Michele A. Paludi and Lisa A. Strayer. What's in an author's name? Differential evaluations of performance as a function of author's name. *Sex Roles*, 12(3-4):353–361, February 1985. ISSN 0360-0025, 1573-2762. doi: 10.1007/BF00287601.
47. Julie R. Posselt. *Inside Graduate Admissions*. Harvard University Press, 2016. ISBN 978-0-674-08869-6.
48. Jeff Colgan. Gender Bias in International Relations Graduate Education? New Evidence from Syllabi. *PS: Political Science & Politics*, 50(02):456–460, April 2017. ISSN 1049-0965, 1537-5935. doi: 10.1017/S1049096516002997.
49. Bart Penders. Ten simple rules for responsible referencing. *PLOS Computational Biology*, 14(4):e1006036, April 2018. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1006036.
50. Jane Lawrence Sumner. The Gender Balance Assessment Tool (GBAT): A Web-Based

Tool for Estimating Gender Balance in Syllabi and Bibliographies. *PS: Political Science & Politics*, 51(02):396–400, April 2018. ISSN 1049-0965, 1537-5935. doi: 10.1017/S1049096517002074.
51. Gabriella Gutiérrez y Muhs, Yolanda F. Niemann, Carmen G. González, and Angela P. Harris, editors. *Presumed incompetent: the intersections of race and class for women in academia*. University Press of Colorado, Boulder, Colo, 2012. ISBN 978-0-87421-869-5 978-0-87421-870-1.
52. The UP Directory: A directory of philosophers from underrepresented groups in philosophy.
53. John Rawls. *A theory of justice*. Belknap Press of Harvard University Press, Cambridge, Mass, rev. ed edition, 1999. ISBN 978-0-674-00077-3 978-0-674-00078-0.
54. Julian Lamont and Christi Favor. Distributive Justice. In *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (Ed.), winter 2017 edition edition, 2017.
55. Serena Olsaretti. *The Idea of Distributive Justice*, volume 1. Oxford University Press, June 2018. doi: 10.1093/oxfordhb/9780199645121.013.38.
56. Iris Marion Young and Danielle S. Allen. *Justice and the politics of difference*. Princeton University Press, Princeton, N.J, paperback reissue edition, 2011. ISBN 978-0-691-15262-2. OCLC: ocn751237488.
57. Sara Ahmed. *On being included: racism and diversity in institutional life*. Duke University Press, Durham ; London, 2012. ISBN 978-0-8223-5221-1 978-0-8223-5236-5.
58. Margaret Urban Walker. *What is reparative justice?* Number 2010 in Aquinas lecture. Marquette University Press, Milwaukee, Wis, 2010. ISBN 978-0-87462-177-8. OCLC: ocn503596039.
59. Elizabeth Anderson. *The imperative of integration*. Princeton University Press, Princeton, N.J, 2010. ISBN 978-0-691-13981-4. OCLC: ocn587248993.
60. Carl T. Bergstrom. Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News*, 68(5):314–316, May 2007. ISSN 2150-6698, 0099-0086. doi: 10.5860/crln.68.5.7804.
61. Csaba Toth, Elizabeth Durham, Murat Kantarcioglu, Yuan Xue, and Bradley Malin. SOEMPI: A Secure Open Enterprise Master Patient Index Software Toolkit for Private Record Linkage. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, 2014:1105–1114, 2014. ISSN 1942-597X.
62. Simon N. Wood. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edition, 2017.